40 LEU	TRP-78	xx
42 ASP	LEU-87	
43 VAL	GLN-90	
53 LEU	ILE-78	xx
56 PHE	VAL-67	xx
67 ASP	PHE-87	
80 LEU	ILE-109	
92 GLY	PHE-106	xx
95 TRP	GLN-101	xx

TABLE III. Tertiary Constraint Lists for Atim

Set of 62 constraints	Set of 50 constraints	Set of 37 constraints	Set of 62 constraints	Set of 50 constraints	Set of 37 constraints
2-228	xx	xx	91-125	xx	87-120
4-37	xx		91-231	xx	90-122
4-206	xx	xx	93-125	xx	
6-123	xx	xx	94-166	xx	
6-89			95-168	xx	
6-162			98-126	xx	xx
7-248	xx	xx	98-145	xx	xx
10-94	xx		105-145	xx	
11-64	xx	xx	105-148	xx	
11-237	xx	xx	109-152	xx	
15-46	xx	xx	112-149	xx	xx
20-49			112-161	xx	xx
23-237		24-54	116-153	xx	121-160
27-59			127-145	xx	
27-241		32-59	127-165	xx	
30-245	xx	xx	128-142		
36-58	xx	xx	128-165	xx	xx
26-248	xx	41-91	130-175	xx	xx
37-89	xx		133-181	xx	
39-123	xx	44-82	142-165	xx	
47-63			142-189	xx	xx
47-87			143-192	xx	
51-86	xx	xx	150-197	xx	
59-245	xx	xx	155-200	xx	xx
60-89			162-208	xx	xx
63-90	xx		165-189	xx	xx
66-79		67-111	165-209	xx	xx
68-114	xx		183-225	xx	xx
79-114	xx		193-205	xx	xx
89-162		82-120	215-244	xx	xx

90-122	xx	xx	230-248	xx	

Results of Monte Carlo Simulated Annealing

The results of stage 2 are compiled in Table IV, below. The numbers of constraints are given next to protein PDB codes.[14] An estimate of the cRMSD from the PDB structure and conformational energy (in dimensionless k_BT units) is given for the last snapshot of each trajectory. The cRMSD is measured between the Cα's of the real structure and the roughly estimated position of the Cα's of the model chain. The latter are obtained according to the following definition:

$R^c_{\alpha i} = (4\mathbf{r}_i + \mathbf{r}_{i-1} + \mathbf{r}_{i+1})/6$, where the sum in the brackets is over the corresponding side chain coordinates of the model chain. The exact agreement of the secondary structure of the predicted fold and the experimental structure was not examined in detail; however, in all runs, it was very close to the target with a small tendency for extension (by one or two residues) of helical fragments in some cases (e.g., the short helix of plastocyanin). The cRMSD and the energy (in dimensionless k_BT units) correspond to the last snapshots of the second simulated thermal annealing runs.

Generally, the predicted structures cluster into two well-defined groups, one of this dominates on the basis of energy, and which is taken to approximate native structure. The remaining, misfolded structures (when observed more than once) were also similar to each other. They represent the topological mirror structure where the chirality of the connections between secondary structural elements (helices and β-strands) is reversed, but the chirality of the secondary structure elements is the same as in the native state, e.g., helices remain right handed. Several interesting observations emerge from the results presented in Table IV, below. First, in the majority of the runs, the native fold is recovered. The accuracy depends on protein size and number of constraints, but only slightly on protein type. Generally, accuracy increases with decreasing protein size. The best accuracy is observed for the 56-residue, B1 domain of protein G,[41] where in most simulations the obtained

structures had cRMSD from native below 3 Å. Interestingly, for the smaller 6pti fragment with a larger number of constraints, the accuracy was systematically somewhat worse. This reflects the effect of protein "regularity." The fold of protein G has a high content of regular secondary structure, while in the 6pti fragment, a substantial fraction of the chain is classified as a loop or coil. The analysis of other cases shows a tendency towards higher accuracy for more regular folds. The accuracy of helical and α/β proteins is greater than for all β-proteins. This is clearly demonstrated on comparison of 1pcy with 2trx. While both proteins are of comparable size, for 2trx with 16 constraints, structures with a cRMSD below 3.5 Å are produced, but for 1pcy with 15 constraints, structures above 5.2 Å result.

TABLE IV. Coordinate cRMSD and Conformational Energy of the Final Structure at the End of the Simulated Thermal Annealing Procedure

Name	Run no.	cRMSD (Å)	Energy
6pti(18)[a]	1	3.3[b]	-321.9
41 res	2	3.8	-313.2
(18-56 fragment)	3	4.1	-302.8
6pti(9/S1)	1	4.1	-336.4
	2	4.2	-345.4
	3	3.6	-318.9
	4	3.8	-385.9
6pti(9/S2)	1	3.8	-331.8
	2	4.3	-320.2
	3	4.0	-341.6
	4	4.4	-353.6
6pti(9/S3)	1	3.4	-303.1
	2	4.0	-318.7
	3	4.8	-324.5
	4	MI[c]	-323.2
	5	MI	-322.5
6pti(9/S4)	1	MI	-319.1
	2	3.8	-312.8
	3	4.0	-320.8
	4	4.2	-280.4
	5	4.1	-302.0
6pti(9/S5)	1	3.9	-370.0
	2	MI	-324.7